JVI Journals.ASM.org

The Avian XPR1 Gammaretrovirus Receptor Is under Positive Selection and Is Disabled in Bird Species in Contact with Virus-Infected Wild Mice

Carrie Martin,[a] Alicia Buckler-White,[a] Kurt Wollenberg,[b] Christine A. Kozak[a]

Laboratory of Molecular Microbiology,[a] Bioinformatics and Computational Biosciences Branch,[b] National Institute of Allergy and Infectious Diseases, Bethesda, Maryland, USA

Xenotropic mouse leukemia viruses (X-MLVs) are broadly infectious for mammals except most of the classical strains of laboratory mice. These gammaretroviruses rely on the XPR1 receptor for entry, and the unique resistance of laboratory mice is due to two mutations in different putative XPR1 extracellular loops. Cells from avian species differ in susceptibility to X-MLVs, and 2 replacement mutations in the virus-resistant chicken XPR1 (K496Q and Q579E) distinguish it from the more permissive duck and quail receptors. These substitutions align with the two mutations that disable the laboratory mouse XPR1. Mutagenesis of the chicken and duck genes confirms that residues at both sites are critical for virus entry. Among 32 avian species, the 2 disabling *XPR1* mutations are found together only in the chicken, an omnivorous, ground-dwelling fowl that was domesticated in India and/or Southeast Asia, which is also where X-MLV-infected house mice evolved. The receptor-disabling mutations are also present separately in 5 additional fowl and raptor species, all of which are native to areas of Asia populated by the virus-infected subspecies *Mus musculus castaneus*. Phylogenetic analysis showed that the avian XPR1 gene is under positive selection at sites implicated in receptor function, suggesting a defensive role for XPR1 in the avian lineage. Contact between bird species and virus-infected mice may thus have favored selection of mouse virus-resistant receptor orthologs in the birds, and our data suggest that similar receptor-disabling mutations were fixed in mammalian and avian species exposed to similar virus challenges.

The transmission of retroviruses to new hosts can result in the emergence of new infectious diseases and can alter the host genomic architecture and gene-regulatory networks, but the factors that determine whether a pathogen can successfully infect a novel host are poorly understood (1). The xenotropic mouse leukemia viruses (X-MLVs) are gammaretroviruses originally isolated from laboratory mice that are unable to infect the cells of these mice (2). These viruses rely on the XPR1 cell surface receptor for entry (3–5), and receptor orthologs from other mammals, like humans and cats, as well as the various species of wild mice, all permit X-MLV entry (6, 7). The resistance of laboratory mouse cells to infection is due to mutations that alter two XPR1 residues, K500E in the putative third extracellular loop (ECL3) and T582Δ in the fourth loop (ECL4) (8).

There is substantial sequence variation in the receptor-determining regions of the mammalian XPR1 orthologs, and some of these receptor variants show altered ability to mediate entry of one or more of the 6 known host range variants in the XP-MLV family of xenotropic/polytropic viruses (6, 9). In addition to X-MLV, this family includes viruses first described as broadly polytropic (P-MLVs) because of their ability to infect mouse cells, as well as cells of other species (10, 11). Other variants, like xenotropic murine leukemia virus-related virus (XMRV) (12) and several wild mouse isolates, have atypical host ranges defined by differences in their abilities to infect cells with variant XPR1 receptors (6, 7). These infectivity differences are due to specific polymorphisms in the receptor and in the viral envelope glycoprotein, and at least 6 XPR1 residues that modulate receptor sensitivity to the various XP-MLVs have been identified (6, 9).

The various XP-MLVs are found as endogenous and infectious viruses in laboratory mouse strains and in the house mouse subspecies. Although the common inbred strains carry multiple copies of both X-MLVs and P-MLVs, these virus subgroups are largely segregated in the three lineages of the house mouse, subspecies with largely nonoverlapping natural ranges in Eurasia (13). *Mus musculus domesticus*, which carries exclusively P-MLVs, is found in western Europe and the Mediterranean. *Mus musculus musculus* populates the area from eastern Europe through northern Asia to the Pacific, and *Mus musculus castaneus* is native to Southeast Asia and India, where *Mus* originated and which is also the site of the house mouse radiation. *M. m. musculus* and *M. m. castaneus* both carry predominantly X-MLVs as germ line copies, and infectious virus has also been isolated from mice of both species. These 3 house mouse subspecies carry 3 different XPR1 variants, none of which corresponds to the disabled laboratory mouse *Xpr1* variant (6). *M. m. domesticus* has a fully permissive XPR1, but the X-MLV-infected subspecies carry two different restrictive *Xpr1* variants; these subspecies are resistant to P-MLVs, but the XPR1 receptor of *M. m. castaneus* is a more efficient X-MLV receptor than that of *M. m. musculus*. These Eurasian house mouse species are all human commensals and were passively transported by humans to the New World, where *M. m. domesticus* is now widespread.

Because XP-MLV-infected mice are distributed globally, because all mammals but the laboratory mouse have X-MLV-sus-

Received 17 May 2013 **Accepted** 1 July 2013
Published ahead of print 10 July 2013
Address correspondence to Christine A. Kozak, ckozak@niaid.nih.gov.
Supplemental material for this article may be found at http://dx.doi.org/10.1128/JVI.01327-13.

doi:10.1128/JVI.01327-13

TABLE 1 Primers used to amplify full-length *XPR1* from duck, quail, and chicken cells and *XPR1* exons from genomic DNA and primers used for mutation

Primer or primer pair	Sequence (5′–3′)		Amplicon size (bp)
	Forward primer	Reverse primer	
a/c	ATGAAGTTCGCCGAGCACCTCTC	GCATGTGTTCACTTCTTTTGCTAGGTCC	2,079
b/d	CGGCGGCAGGATGAAGTTCG	GCTGCATGTGTTCACTTCTTTTGCTAGGTCC	2,092
e/f	CGAGTGTTCACTGCTCCCTTCC	CTGTATTTGCCAGAAGCCCATC	163
g/h	CATCCCTGCTTGGTTGCG	GGAGATGAGGAACTTCAGGGATCAGAGC	891
g/j	CATCCCTGCTTGGTTGCG	GTCGTTTCCAGGGAGAT	902
k/l	GCTAAAAACCACAGCGACACC	GTGGGTAGACAATTCCTTCTCGAAGAAAGG	163
m1/o1	GCATACTACTATTGTGCC	CTGAAAACCTCAAGTGGGGCGAAAACAG	140
m2/o2	GCGTACTACTATTGTGCC	CTGAAAACCTCAAGCGGGGCAAAAACAG	140
m3/o3	GCTTACTACTACTGTGCC	CTGAAAACCTCAAGTGGGGCAAACACAG	140
n/o3	CCTGCGCTTTGCGTGGACC	CTGAAAACCTCAAGTGGGGCAAACACAG	105
Primer for mutation[a]			
E579Q Chicken	CCAGATCTCCCTCACTTCCATGCAGATCTTCCCATACG		
Q496K Chicken	GCAGCCCTCTACAGCACTCACAAAGCTAAAAACCAC AGCGACACC		
Q579E Duck	CCAGATCTCCCTCACCTCCATGGAAATCTTCCCGTATGCTGG		
K496Q Duck	GCCCTCTACAGCACTCACCAAGCTAAAAACCACAGCGAC		
K496E Duck	GCAGCCCTCTACAGCACTCACGAAGCTAAAAACCACAGCGAC		

[a] Reverse primers were reverse complements.

ceptible XPR1 receptors, and because XPR1 orthologs are found in all eukaryotes, we looked for naturally occurring restrictive receptors in a related set of species outside the mammalian lineage. We focused on avian species for several reasons. First, early studies on X-MLVs suggested that avian cells are variably susceptible to these viruses (14, 15). Second, some birds and mice share habitats, sometimes with a predator-prey relationship. Third, RNA viruses more readily jump species than other pathogens, and there is documented evidence of epizoonotic retroviral transmission from mammals to birds (16, 17).

MATERIALS AND METHODS

Cells, viruses, pseudoviruses, and infection assays. CAST-X is an X-MLV isolated from the spleen of a CAST/EiJ mouse (18). XMRV (12) was kindly provided by R. Silverman (Cleveland Clinic, Cleveland, OH). Cz524 MLV was isolated from the spleen of a CZECHII/EiJ mouse 2 months after inoculation with Moloney ecotropic MLV. CasE#1 (19), FrMCF P-MLV, MoMCF P-MLV, AKR6 X-MLV, and NZB-IU-6 X-MLV were originally obtained from J. Hartley (NIAID, Bethesda, MD). LacZ pseudotypes were generated for these viruses by infection of the packaging cell line GP2-293 (Clontech, Mountain View, CA) that had been transfected with pCL-MFG-LacZ (Imgenex, San Diego, CA), along with pM-SCVpuro (Clontech), by J. Silver (NIAID, Bethesda, MD). Filtered media from the virus-infected cultures contained a mixture of infectious virus and LacZ pseudovirus.

Susceptibility to replication-competent virus was quantitated by infecting cells with dilutions of virus stocks in the presence of Polybrene (4 to 8 μg/ml; Aldrich, Milwaukee, WI). After 4 days, cultures were UV irradiated and overlaid with mink $S^+ L^-$ cells (20). Foci of transformed $S^+ L^-$ cells that mark virus-producing infected cells were counted 5 to 7 days later, and titers were expressed as focus-forming units/200 μl.

Virus susceptibility was also assessed in a single-round assay using XP-MLV pseudoviruses to infect *Mus dunni* cells (21), Chinese hamster E36 cells (22), ferret MA139 cells obtained from J. Hartley, duck cells (ATCC CCL-141), quail cells (ATCC CRL-1708), and primary cell cultures of chicken embryo fibroblasts made from chick embryos. The cells were infected with 10-fold dilutions of pseudotype stocks in the presence of Polybrene. One day after infection, the cells were fixed with 0.4% glutaraldehyde and assayed for β-galactosidase activity using as the substrate 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) (2 mg/ml; ICN Biomedicals, Aurora, Ohio). Infectious titers were expressed as the number of blue cells per 50 μl. *P* values were determined from a 2-tailed Student's *t* test using GraphPad Prism 6.

Avian *XPR1* sequencing. DNA and RNA were isolated from tissue samples and cultured cells from various bird species (see Table S1 in the supplemental material). Tissue samples for DNA extraction were provided by J. Dean (Smithsonian National Museum of Natural History, Washington, DC) and S. Rasheed (University of Southern California, Los Angeles, CA). Additional DNA samples were provided by R. Kimball (University of Florida, Gainesville, FL) and G. Proudfoot (Vassar College, Poughkeepsie, NY).

Total RNA was extracted from duck, chicken, and quail cells using TriReagent (Molecular Research Center, Cincinnati, OH). cDNA was synthesized using the SuperScript II reverse transcriptase kit (Invitrogen, Carlsbad, CA). Full-length *XPR1* was amplified using primers designed from the sequenced chicken genome (23) (Table 1) and cloned into the expression vector pcDNA3.1/V5-His TOPO (Invitrogen). Our cloned chicken *XPR1* differs from the *Gallus* sequence at one site, A8S.

Segments of DNA including *XPR1* exons 10 to 13 were amplified from genomic DNAs using primers designed from coding or intron sequences based on the sequenced chicken genome or on an alignment of the sequenced chicken, turkey, and zebra finch genomes (Table 1). These PCR primers, as well as others not listed, failed to produce amplicons for some segments in some species, as noted (see Text S1 in the supplemental material). PCR products were cloned into pCR2.1-TOPO (Invitrogen, Carlsbad, CA) for sequencing. The sequenced segments varied from 407 bp (lark) to 1,468 bp (ostrich and rhea).

Generation of *XPR1* mutants. Six mutant variants of the duck and chicken *XPR1* genes were generated by mutagenesis PCR using the QuikChange II site-directed mutagenesis kit (Stratagene, La Jolla, CA). The primers used for mutagenesis are listed in Table 1. A chimeric *XPR1* backbone was generated by replacing a segment of the chicken *XPR1* with a 590-bp EcoRI-HpaI segment of the duck gene containing the N-terminal end of ECL3. The PCR fragment used for the replacement was generated with the forward and reverse primers 5′CCTGATAGAATTCCTCT TTCTCCTGG and 5′-CCAGAGTTAACCAGATGAAAGGC. Q496K and E579Q were introduced into this chimera. All mutants were confirmed by sequencing.

The recombinant plasmids were transfected into E36 Chinese hamster cells with Fugene6 (Roche, Indianapolis, IN). E36 has a functional XPR1 receptor, but virus entry through the endogenous receptor is blocked by

glycosylation; these cells show marginal susceptibility to X-MLVs and resistance to all other XP-MLVs (9, 24). At least 2 pools of stable transfectants were selected for each mutant with Geneticin (830 μg/ml; Invitrogen).

Expression of *XPR1* was confirmed by Western analysis. Proteins were extracted from transfected cells with M-PER Mammalian Protein Extraction Reagent (Pierce, Rockford, IL). The expression vector used for *XPR1* inserts a V5 epitope at the C terminus; *XPR1* expression was detected in Western blots using anti-V5 antibody (Invitrogen), followed by goat anti-mouse IgG conjugated with horseradish peroxidase (HRP) (Invitrogen). The membrane was then stripped and incubated with mouse anti-α-tubulin (Sigma, St. Louis, MO) and goat anti-mouse IgG conjugated with HRP (Invitrogen).

Selection analysis. DNA sequences were aligned using CLUSTALW2 and improved manually. Two phylogenies were produced, one for a set of 23 sequences that included segments of *XPR1* exons 10 to 13 and one for a set of 31 sequences representing only the receptor-determining regions in exons 11 and 13. Both phylogenies were cladograms that were manually constructed to correspond to the consensus *Aves* phylogeny (25–28).

The codeml program of the PAML4 package (29) was used for maximum-likelihood analysis of positive selection on codons. This analysis was performed using two different codon frequency models: the F3x4 model (with codon frequencies estimated from the nucleotide frequencies at each codon site in the data) and the F61 model (with codon frequencies of each of the 61 nonstop codons calculated from the data). The neutral-selection/negative-selection model of codon selection (M7) was compared to the corresponding positive-selection model (M8, which includes a category for ratios of nonsynonymous to synonymous evolutionary rates [*dN*/*dS* ratios] of >1). The significance of this additional codon selection category was assessed using a likelihood ratio test of the phylogeny likelihoods under the neutral- and positive-selection models. The significance of the test statistics was calculated using a chi-squared distribution with 2 degrees of freedom. The Bayes empirical Bayes algorithm was used to calculate the posterior probability of individual codons experiencing *dN*/*dS* ratios of >1.

Nucleotide sequence accession numbers. The GenBank accession numbers for the sequenced segments are JX294430 to JX294458.

RESULTS

Susceptibility of avian cells to XP-MLVs. Duck and chicken cells were assessed for susceptibility to infectious XP-MLVs. Cells were infected with serial dilutions of two X-MLVs, NZB-9-1 and the wild mouse-derived CAST-X, and two P-MLVs, FrMCF and MoMCF (Fig. 1A), and foci of infected cells were detected using the mink $S^+ L^-$ indicator cell line (20). Duck cells were efficiently infected with both X-MLVs, as were cells of the Asian mouse *M. dunni*, which carries the fully permissive *Mus Xpr1*^{sxv} receptor variant. Duck cells were considerably less susceptible to the P-MLVs, showing a 1,000- to 10,000-fold titer reduction compared to that of *M. dunni* cells (Fig. 1A). In contrast, chicken cells were completely resistant to 3 of the 4 viruses and showed poor infectivity by CAST-X X-MLV.

We then tested chicken, duck, and quail cells for virus susceptibility in a single-round infection assay using XP-MLV pseudoviruses (Fig. 1B). This virus set included 3 X-MLVs, 1 P-MLV, and 2 wild-mouse isolates with novel host ranges, CasE#1 and Cz524 (9). These 6 viruses were chosen because they have distinctively different host ranges on nonrodent mammalian cells and on cells expressing mutated variants of *Mus Xpr1* (6, 9). The viruses rely on different but overlapping sets of determinants on *Mus Xpr1* for entry. The pattern of pseudovirus infection for each cell line was consistent with the results obtained with replicating virus (Fig. 1A). *M. dunni* cells are efficiently infected by all pseudoviruses, but

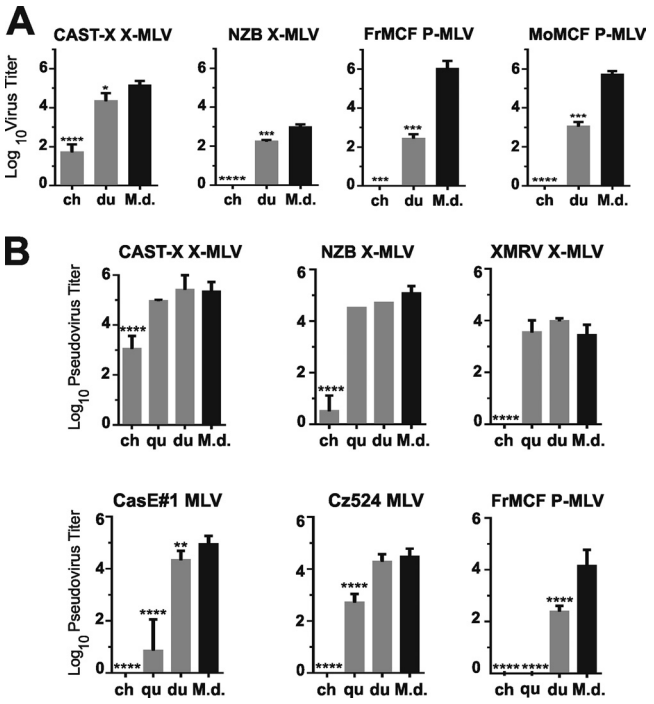

FIG 1 Infectivity of XP-MLVs in avian cells. (A) Susceptibility of cells to infectious XP-MLVs. Cells were infected with virus dilutions, UV irradiated 4 days later, and overlaid with mink $S^+ L^-$ cells. Infected, virus-producing avian cells induced transformed foci in the indicator cells. The titers represent focus-forming units in 200 μl plus standard errors of the mean (SEM). (B) Susceptibility to LacZ pseudovirus. Cells were stained to detect the reporter gene 24 h after infection with pseudovirus dilutions. The titers represent blue cells in 50 μl of virus stock and are presented as the means of 3 to 5 tests plus SEM. In some cases, there were no positive cells after infection with undiluted virus stock. The asterisks indicate significant *P* values determined from 3 or 4 independent tests using Student's *t* test (*, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$; ****, $P \leq 0.0001$); infectivity of the avian cells was compared to the permissive *M. dunni* (M.d.) mouse line. ch, chicken; qu, quail; du, duck.

the 3 avian cells show three different susceptibility patterns (Fig. 1B). Chicken cells are inefficiently infected by X-MLVs and are resistant to other XP-MLV pseudoviruses. The quail and duck cells were efficiently infected by X-MLVs, and both showed reduced susceptibility to P-MLV, but quail cells are also poorly susceptible to the 2 wild-mouse viruses. These results indicate that infection with XP-MLVs is restricted in chickens at an early stage in the virus life cycle and that there is no postentry restriction of X-MLV replication in duck cells.

Functional evaluation and mutational analysis of avian XPR1 receptors. To determine if susceptibility differences are receptor mediated, we sequenced the XPR1 receptor genes from the 3 avian cell lines (duck, quail, and chicken). These genes are 94 to 97% identical to one another (see Fig. S1 in the supplemental material) and are 92 to 99% identical to the XPR1 genes from the sequenced turkey and zebra finch genomes (NC_015020 and NW_002198997). The avian XPR1 protein is ~83% identical to the permissive *M. dunni* XPR1, although identity drops to 38% in the receptor-determining ECL4. Comparison of the 3 avian genes identified 27 replacement mutations (see Fig. S1 in the supplemental material), but only three substitutions distinguish the virus-resistant chicken XPR1 from both of the more permissive duck and quail receptors. Two of the three substitutions are at sites

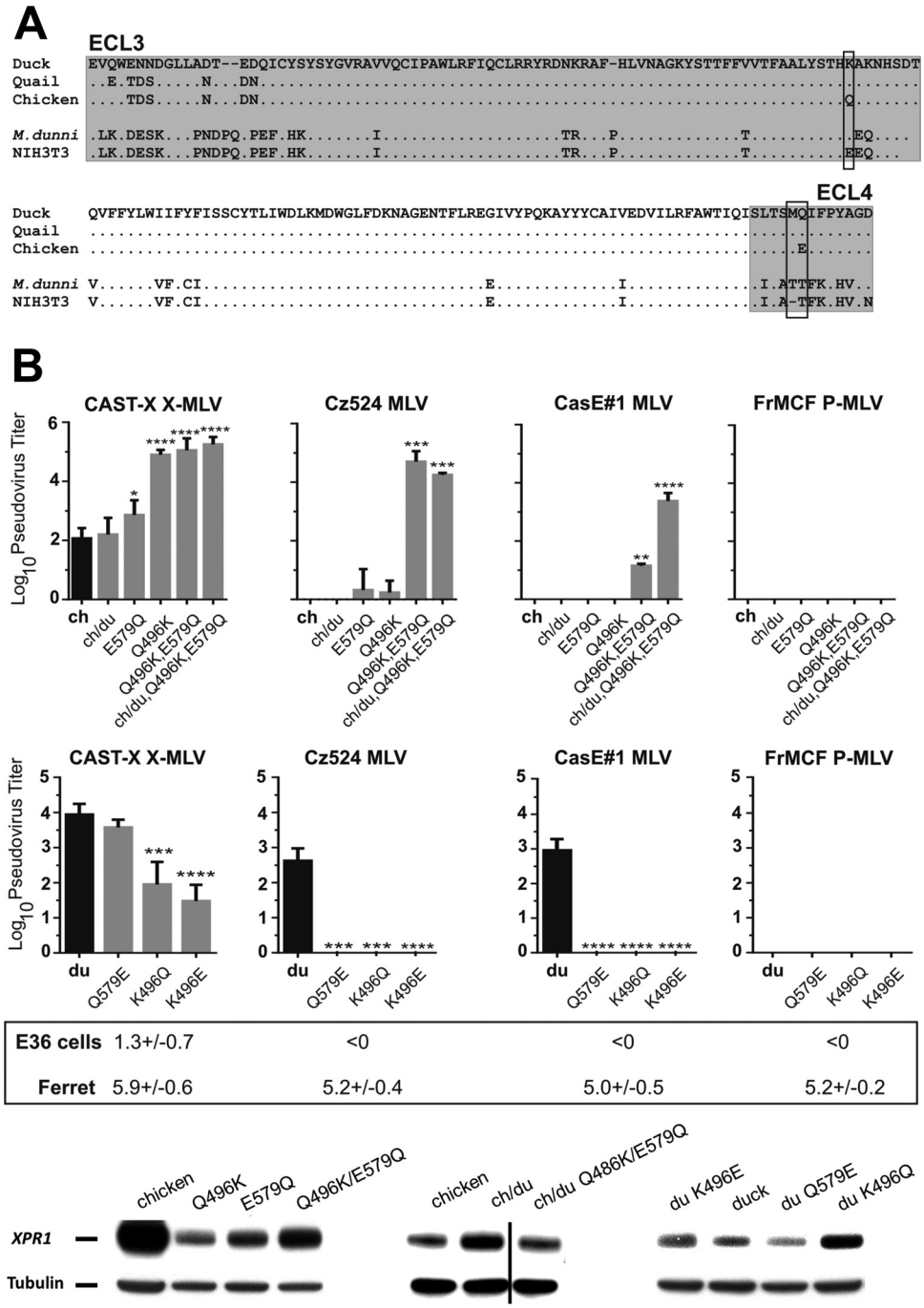

FIG 2 Identification and evaluation of XPR1 replacement substitutions for possible roles in receptor function. (A) Alignment of the segment of avian and mouse XPR1 encoding putative extracellular loops ECL3 and ECL4. The shaded areas represent the ECLs, dots represent identical residues, and the boxes identify critical sites that distinguish XPR1 in virus-susceptible and -resistant mouse and avian cells. The NIH 3T3 mouse ECL4 deletion removes one of two adjacent T residues. (B) Susceptibility of hamster E36 cells expressing chicken, duck, and mutant XPR1 variants to LacZ pseudoviruses of 4 XP-MLV host range variants. The K496E substitution introduced into the duck construct was identified in avian species other than chicken. The box shows titers for untransfected E36 cells and fully permissive ferret cells used as a positive control. The asterisks identify significant *P* values determined from 3 or 4 independent tests using Student's *t* test (*, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$; ****, $P \leq 0.0001$); titers for each of the chicken or duck mutants were compared to their parental types. At the bottom is a Western blot showing expression of V5-tagged XPR1s in transfected E36 cells; the vertical black line indicates a deleted lane.

that are homologous to the two substitutions that disable the receptor function of laboratory mouse XPR1, namely, K496Q in ECL3 (K500E in mouse) and Q579E in ECL4 (T582Δ in mouse, which deletes one of two adjacent Thr residues) (Fig. 2A; see Fig. S1 in the supplemental material). The third substitution that differentiates chicken XPR1 from the other avian gene products, E261D, is in ECL1, a region that has not been implicated in XPR1 receptor function. Previous analysis of *Mus Xpr1* variants identi-

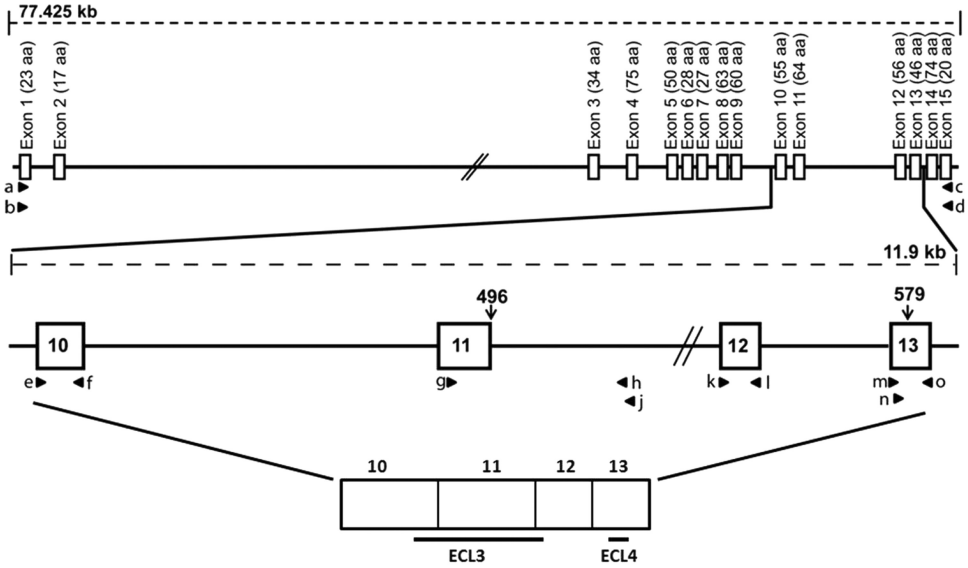

FIG 3 Structure of the avian *XPR1* gene with primers used to amplify exons 10 to 13 from genomic DNA and the full-length gene from cDNA. Exon and intron primers were designed from the chicken gene or based on an alignment of the genes from the sequenced chicken, turkey, and zebra finch genomes (23, 30, 31) (Table 1). At the bottom is the protein sequence showing the positions of the two putative extracellular loops within the 4 sequenced exons. aa, amino acids.

fied 6 sites in addition to 500 and 582 that influence entry by the various XP-MLVs (6, 8, 9, 18); all of these sites are found in the C-terminal end of ECL3 or in ECL4, and all 6 of the sites are invariant in the 3 bird genes.

The chicken and duck genes were cloned into an expression vector and transfected into virus-resistant Chinese hamster E36 cells (Fig. 2B). Transfectants were infected with pseudoviruses representative of 4 distinct host range types: an X-MLV, a P-MLV, and the two wild-mouse isolates. The transfected cells largely recapitulated the susceptibility patterns of the chicken and duck cells. Transfectants expressing the chicken gene showed very poor susceptibility to X-MLVs and were resistant to the other XP-MLV pseudoviruses. E36 cells expressing the duck *XPR1* were highly susceptible to most XP-MLVs, although susceptibility to P-MLVs, which is marginal in duck cells (Fig. 1), was not detected in these transfectants.

We introduced reciprocal mutations into the cloned chicken and duck genes at the two sites implicated by sequence comparisons (496 and 579) and expressed these constructs in E36 cells (Fig. 2B). Mutations at either of the sites in the defective laboratory mouse *Xpr1* gene generate functional X-MLV receptors (8), and improvement of receptor function was also noted for the chicken gene following introduction of these separate mutations, Q496K ($P < 0.0001$) and E579Q ($P = 0.0406$). The reciprocal changes in the duck XPR1, K496Q and Q579E, each reduced receptor function (Fig. 2B). The two mutations were not, however, equivalent; both receptors with the permissive K496 residue showed better receptor function than those with Q579. The introduction of both mutations into the chicken gene produced highly functional receptors, suggesting that these residues cooperate to produce an efficient receptor.

The chicken and duck genes also differ substantially in sequence at the N-terminal end of ECL3, residues 423 to 434 (Fig. 2A). This region of *XPR1* was shown to be under positive selection in the rodent gene (6) but has not been evaluated for a role in receptor function. Therefore, we created a chimeric avian gene (termed ch/du) containing duck ECL3 codons 419 to 474 on the chicken *XPR1* backbone. This substitution alone did not affect receptor function (Fig. 2B), but the addition of the Q496K and E579Q mutations to this chimeric *XPR1* produced a better CasE#1 receptor than the chicken *XPR1* with only 2 mutations ($P = 0.0004$) (Fig. 2B). This suggests that residues in the N-terminal end of ECL3 can modulate receptor function.

To determine if other XP-MLVs display similar infectivity patterns and show additive effects with the double mutation, we used pseudoviruses for three additional X-MLVs (NZB-IU-6, AKR6, and XMRV) and an additional P-MLV (MoMCF). These viruses have sequence differences in the receptor-determining VRA region of the viral *env* and show infectivity differences in mammalian species (6, 19). The three X-MLVs showed infectivity patterns in the transfected cells that resembled each other and CAST-X, and MoMCF, like FrMCF P-MLV, failed to infect any of the transfected cells (data not shown). The additive effects observed for the two chicken constructs with double mutations are thus restricted to the wild-mouse viruses, CasE#1 and Cz524.

Comparative analysis of avian *XPR1* genes. To determine the species and geographic distributions of these disabling mutations in birds, we sequenced receptor-determining segments of *XPR1* in 30 avian species trapped on 5 continents (Fig. 3; see Table S1 and Text S1 in the supplemental material) and extracted the corresponding segments from the sequenced genomes of the turkey and zebra finch (30, 31). Overrepresented in this data set are Galliformes species related to domestic chickens and species of avian raptors. The mouse and chicken XPR1 proteins are encoded by 15 exons spanning 142.9 kb in the mouse genome and 77.4 kb in the sequenced chicken genome (Fig. 3). The putative 85-residue receptor-determining ECL3 is contained in 3 exons, and the shorter 13-residue ECL4 is encoded by exon 13.

The sequenced segments include the entire ECL4, the variable 5′ end and 3′ end of ECL3, and the 5′ end of intron 11 (see Text S1

		ECL3 496	ECL4 574 - 586
Birds	Duck, quail, 17 species	K	SLTSMQIFPYAGD
	Chicken	Q	.....E.....
	Grey junglefowl	E	.....
	Green junglefowl	Q	.....
	Sri Lankan junglefowl	E	.....
	Philippine hawk-owl	E	.....
	Rhea	E	.....
	Chinese partridge	.	.....E.....
	Nuthatch	.	.....A.
	Zebra finch	.	.....A.
	Pelican	.	.....H...
	Heron	.	.....S.
		500	578 - 590
Mammals	*M. dunni*	K	SITATTFKPHVGD
	Laboratory strains	E	....-.....N
	M. m.molossinus	.	.-...K.....
	M.pahari	.	...V.....
	M. m. castaneus	.	....-----....
	M. tenellus	.	.....R.....
	A.niloticus	.	.....L.....
	H.univitatus	.	.M...A...R...
	U. ruddi	.	...V.SLQ.N...
	L.sikapusi	.	...H.ALQ.....
	Gerbil	.	...V...Q.....
	Chinese hamster	.	.....A.Q.....
	Guinea pig	.	...TMPTL..S..
	Kangaroo rat	.	...S..KLSN...
	Deer mouse		...TMPTL..S..
	Primate, rabbit, elephant	.	...S..LL..S..
	Pig	.	...T..SL..S..
	Mink, cat, horse	.	...SM.LL..S..
	Dog	.	.S.SM.SL..S..
	Cow,buffalo,goat	.	...SM.SL..S..
	Bat	.	.V.S..LM..T..
	Armadillo	.	...SM.LL.....
	Platypus	D	.L.TMDIF..A..

FIG 4 Variation in the deduced protein sequences of the receptor-determining regions of avian and mammalian *XPR1*. Shown are avian and mammalian variants that differ, respectively, from the permissive duck and *M. dunni* genes. Residue site numbers are given, and the critical sites for receptor function are boxed.

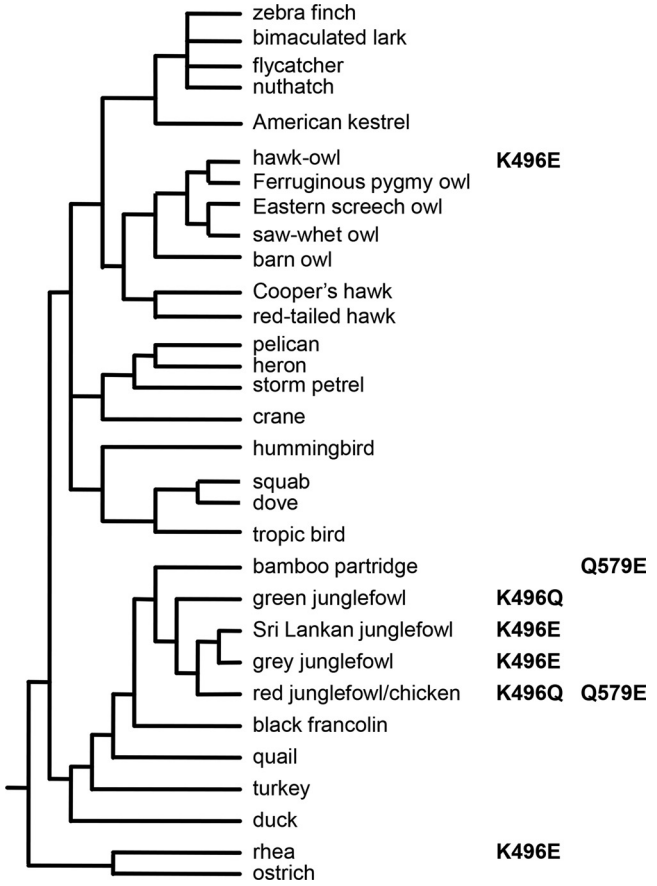

FIG 5 Distribution of disabling *XPR1* mutations in avian species. (The phylogenetic tree is adapted from references 25 to 28.)

in the supplemental material). The sequenced exon segments show 93 to 99% identity across all avian species. The avian *XPR1* ECL4 is much less variable than the comparable segment of the mammalian gene (2). Rare replacement mutations alter only 3 ECL4 codons in bird species (Fig. 4), whereas for mammalian XPR1 genes, there are substitutions affecting all but 3 ECL4 residues. The 3 conserved mammalian ECL4 sites do not contribute to the virus attachment site in *Mus Xpr1* (6). Although there is otherwise limited protein sequence identity between the ECL4 domains of birds and mammals (38% identity with mouse), two of the 3 ECL4 residues conserved in mice are also conserved in bird XPR1 (S574 and T576), suggesting they have been subjected to purifying selection.

Of the two receptor-inactivating mutations identified in the chicken, one, the ECL4 substitution Q579E, was found in one other Galliformes species, the Chinese bamboo partridge, *Bambusicola thoracica* (Fig. 5; see Text S1 in the supplemental material). No other avian species carried this mutation or any other nonsynonymous mutation at the site.

The substitution with the greater impact on receptor function, K496Q (Fig. 2B), was identified only in chicken and one other *Gallus* species, the green junglefowl, *Gallus varius* (Fig. 5). The 2 mutations that disable the chicken gene thus either predate domestication or were introduced into the wild through crossbreeding with chickens, a distinct possibility, as chickens and green junglefowl are interfertile (32) and are routinely interbred by the Javanese, producing a hybrid termed the Bekisar.

The sequence analysis also identified a novel substitution altering one of these two critical sites, K496E, in two *Gallus* species, the gray and Sri Lankan junglefowls, *Gallus sonnerati* and *Gallus lafayetti* (Fig. 5). In *Mus*, the homologous ECL4 substitution (K500E) helps disable the laboratory mouse XPR1 (8). To determine the impact of this mutation on avian receptor function, we introduced K496E into the duck clone; it results in significant impairment of receptor function (Fig. 2B). For the other Galliformes species examined, the XPR1 gene of the black francolin (*Francolinus francolinus*), turkey (*Meleagris gallopavo*), and quail (*Coturnix japonica*), resemble the permissive duck gene at both critical sites (Fig. 5).

With two exceptions, 23 additional species from 17 different avian families were found to carry the permissive residues at positions 496 and 579 (Fig. 5; see Text S1 in the supplemental material). One of the disabling mutations, K496E, was also identified in one raptor, the Philippine hawk-owl, *Ninox philippensis*, but was not found in 8 other raptors tested, which included various owls, hawks, and a falcon. Although the same mutation, K496E, was also identified in the South American ratite, the rhea, these results

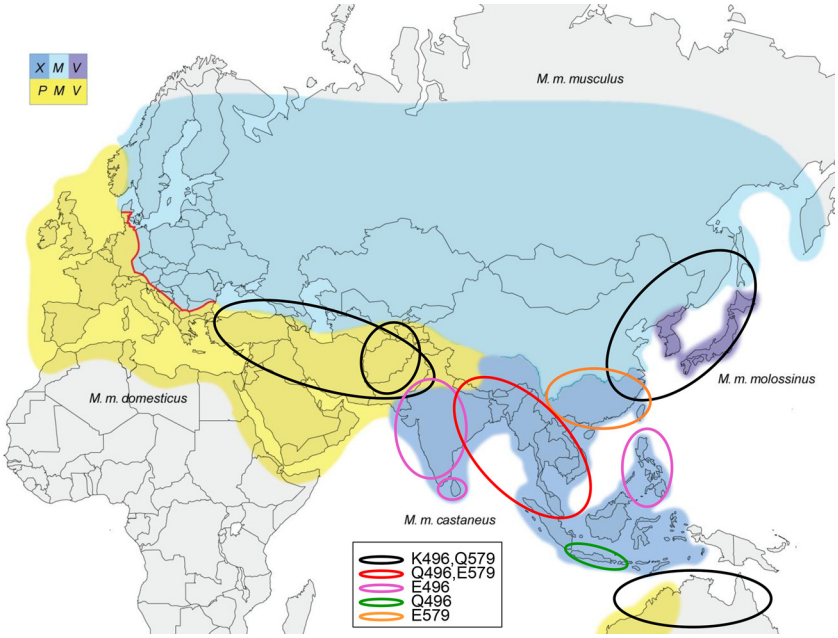

FIG 6 Geographic distribution of avian species and house mouse subspecies. Areas populated by the 3 mouse lineages carrying endogenous X-MLVs (*Xmvs*) are colored in shades of blue, while yellow marks the range of mice with endogenous P-MLVs (*Pmvs*). The ovals of various colors represent the approximate ranges of avian species carrying restrictive *XPR1* mutations, and the black ovals represent ranges of species with permissive residues. Red, red junglefowl (*Gallus gallus*); pink, gray junglefowl (*G. sonnerati*), Sri Lankan junglefowl (*G. lafayetti*), and Philippine hawk-owl (*N. philippensis*); green, green junglefowl (*G. varius*); orange, bamboo partridge (*B. thoracica*); black, lark (*Melanocorypha bimaculata*), quail (*C. japonica*), francolin (*F. francolinus*), and barn owl (*Tylo alba*). The red line is the hybrid zone separating *M. musculus* subspecies in Europe.

indicate that the species that carry the mutations that disable receptor function are largely native to a defined geographic area that encompasses India and Southeast Asia (Fig. 6).

Positive selection of avian *XPR1*. We used the PAML4 suite of programs (29) to test for evidence of adaptive evolution and to identify possible sites of positive selection marked by an excess of nonsynonymous mutations in the sequenced segments of avian *XPR1*. Positive selection is a hallmark of genes involved in genetic conflicts and often identifies genes that serve a defensive function. Evidence of involvement in an evolutionary "arms race" has been found for other genes capable of restricting MLVs, like primate *TRIM5*α (33) and the mouse genes *Fv1* (34) and *Apobec3* (35). There is also evidence that rodent *Xpr1* has been under positive selection at sites coding for ECL3 and ECL4 that have been implicated in receptor function (6). Rodent *Xpr1* variation is coincident with exposure to MLVs; 4 of the 5 polymorphic variants of *Mus Xpr1* restrict different subsets of XP-MLVs, and all 4 variants evolved in Southeast Asia, where mice were first exposed to XP-MLV infection and acquired MLV endogenous retroviruses (ERVs) (7).

Avian sequences of receptor-determining regions of exons 11 and 13 for 31 species and exons 10 to 13 for 23 species were used in this analysis. The phylogeny used was based on the accepted phylogenetic tree for *Aves* (25–28) (Fig. 5). Likelihood ratio tests indicate that *XPR1* has experienced positive selection, and this was the case for both codon frequency models tested (Fig. 7A; see Table S2 in the supplemental material). The Bayes empirical Bayes calculation of posterior probabilities in PAML4 identified one *XPR1* codon position, 496, as being under positive selection in both data sets (>95% posterior probability) (Fig. 7B; see Table S2 and Text S1 in the supplemental material). A second codon in the variable region at the N terminus of ECL3, 423, was also identified as being under positive selection in the 4-exon data set (posterior probability of >0.99). Thus, positive selection affects one codon critical for XPR1 receptor function (496) and a second codon in a region of ECL3 implicated in the modulation of receptor function (Fig. 2B).

Because one of the disabling mutations was found in the rhea, a species with no known exposure to infectious XP-MLVs, we reran the analysis without this species (see Table S3 in the supplemental material), and we also ran the analysis without the passerines and water birds (not shown). Significance levels of selection were reduced, but both sites remained under positive selection in at least one codon frequency model tested.

The analysis indicates that the evolution of this avian gene has been driven by genetic conflicts centered on codons implicated in virus restriction, and this is consistent with an antiviral function for the gene in birds.

DISCUSSION

There are many documented examples of the transspecies transmission of retroviruses (36). Such naturally occurring transmissions can produce disease in the new hosts (HIV-1), and in some cases the introduced viruses can be endogenized (koala endogenous retrovirus [KoRV]) (37). The fact that birds have acquired retroviruses from mammalian sources has been established by the observed phylogenetic discontinuities between birds and the endogenous viruses they carry (16, 17). The results presented here suggest that interorder transmission of X-MLVs from mice to birds may have driven fixation of resistant alleles in the avian ortholog of the X-MLV XPR1 receptor, a phenomenon also doc-

A

Exons 11 and 13

	M7-M8	Tree length	dN/dS
F3x4	p = 0.01	1.67283	3.84
F61	p = 0.02	1.71391	2.75

Exons 10-13

	M7-M8	Tree length	dN/dS
F3x4	p < 0.0001	1.24815	3.37
F61	p < 0.0001	1.25117	3.02

B

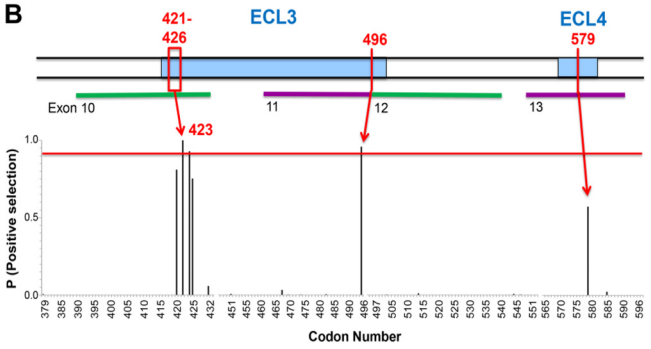

FIG 7 Avian *XPR1* sites under positive selection. (A) Likelihood ratio tests (LRTs) were used to test for positive selection. A neutral model (M7) was compared with a selection model (M8) using two different models of codon frequency (F3x4 and F61). *P* values of <0.05 provide evidence of positive selection, and *P* values of <0.01 indicate strong positive selection. The tree length is the average number of substitutions per codon along all branches. *dN*/*dS* ratios are the mean values estimated by PAML4 for the category *dN*/*dS* ratios of >1. (B) (Top) Diagram of a portion of the XPR1 gene showing the locations of ECL3 and ECL4 and the locations of the key codons involved in receptor function. The four alternating green and purple horizontal lines indicate the positions of the sequenced segments of exons 10 to 13. (Bottom) Graph showing the posterior probability of positive selection at each codon based on an analysis of 22 sequences using codon frequency model F3x4 and selection model 8. The horizontal red line marks a posterior probability greater than 0.95, and the 3 arrows identify the 2 codons over this threshold (and therefore under selection) and critical residue 579.

umented for the same virus-receptor combination in various virus-infected inbred and wild mice (6, 7).

The transspecies transmission of infectious agents is influenced by various geographical, ecological, behavioral, and genetic factors. Among the ecological factors that increase the likelihood of cross-species infection are shared habitat and intimate contact, and domesticated chickens are ground-dwelling birds that not only live in close contact with house mice, but are effective mousers that are omnivorous and can and do eat mice (38) (see Text S2 in the supplemental material). Protection from infectious agents is particularly important in the high-density populations typical of agricultural settings, and while there are many host factors that restrict retrovirus replication, birds lack many of the retrovirus restriction factors found in mice and other mammals, such as APOBEC3, tetherin/BST-2, and Fv1. As shown here, X-MLVs can replicate efficiently in some avian cells, suggesting that birds exposed to virus would benefit from protective mechanisms, like restrictive receptors.

The two inactivating mutations identified in the chicken *XPR1* gene were found together only in chickens, which were domesticated 4,000 to 8,000 years ago in the Indus Valley and/or Southeast Asia (39–41). Darwin originally proposed the red junglefowl, native to these areas (Fig. 6), as the progenitor of chickens (42), although recent studies suggest additional small contributions to the chicken genome from other junglefowl species (43). The domesticated chicken was subsequently introduced to the West through trade between the city states of the Indus Valley and the Middle East and may have been introduced to the New World by Polynesians (38). The Indian subcontinent is also where the 3 lineages of house mice carrying the various XP-MLVs originated before dispersing to their current Eurasian ranges and beyond (44, 45) (Fig. 6). These are *M. m. domesticus* (western Europe and the Mediterranean), which carries endogenous P-MLVs, and *M. m. castaneus* (Southeast Asia) and *M. m. musculus* (central Europe to China), both of which carry mainly X-MLVs (13). These commensal or house mice differ from their *Mus* progenitors by their dependence on humans; the animals live in close contact with humans in manmade structures, while their aboriginal progenitors inhabit various niches in the wild (46). Thus, red junglefowl and house mice both evolved in the same geographic area, and omnivorous chickens were domesticated in human habitations infested with X-MLV-infected *M. m. castaneus* house mice.

Four of 8 additional fowl and 1 of 9 raptors were found to carry 1 of the 3 XPR1 receptor-disabling mutations (K496Q, K496E, or Q579E). All 5 of these species are native to south central or Southeast Asia and thus, like *G. gallus*, have ranges that overlap the range of *M. m. castaneus* (Fig. 6) (45, 47). None of the 3 *XPR1* mutations was found in 11 other fowl and raptor species native to the Middle East, Australia, or the Americas, areas populated largely by *M. m. domesticus*, or in the case of quail, by *M. m. musculus* and *Mus musculus molossinus*, which is a natural hybrid of *M. m. musculus* and *M. m. castaneus* (48). Our analysis included 14 species that are neither fowl nor raptors, such as water birds, passerines, and ratites, representing 100 million years of evolution. Only 1 of these 14 species, the South American rhea, carries mutations at either of the two key sites. The habitat of this bird is thus outside the *M. m. castaneus* range of Southeast Asia, and although *M. m. castaneus* mice have been found in the Americas, in California (13, 49), the few characterized South American house mice, all from Peru, have been typed as *M. m. domesticus* (50). The presence of this mutation in the rhea could thus result from neutral processes or from an unknown nonneutral pressure. Removing the rhea sequence from the PAML4 analysis for positive selection reduces but does not eliminate evidence for positive selection.

Among wild mouse populations, X-MLVs are found in *M. m. molossinus* and *M. m. musculus*, as well as *M. m. castaneus*. These mice all carry endogenous X-MLVs, and infectious virus has been isolated from all 3 subspecies (51–53). This raises the question of why receptor mutations were not found in birds from other parts of Eurasia where these other X-MLV-infected mouse subspecies are found. There are several possible explanations. First, our collection of 32 species did not include raptors likely to have intimate contact with these other Eurasian mouse subspecies. Second, as shown for inbred laboratory strains, mice can vary significantly in their ability to produce infectious X-MLV, with some strains producing high levels of X-MLV throughout life (54, 55). We have no information, however, on the level of endogenous virus production in the 3 X-MLV-positive wild mouse populations, but we do know that the *Xpr1* receptor allele of *M. m. castaneus* ($Xpr1^c$) is a more efficient receptor for X-MLVs than the *Xpr1* variant shared by *M. m. musculus* and *M. m. molossinus* ($Xpr1^m$) (6), and this difference could certainly influence endogenous virus levels.

There is no evidence for endogenous retroviruses closely re-

sembling MLVs in birds. Although one of the studies demonstrating phylogenetic discontinuities between bird species and their endogenous retroviruses identified those avian ERVs using degenerate primers designed from MLVs, the amplified ERVs cluster with mammalian retroviruses but are not MLVs (17). The sequenced *G. gallus* genome lacks recognizable MLV-like ERVs (56). While the presence of MLVs in birds would lend support to a long-term evolutionary arms race, the absence of such sequences is not surprising. After all, endogenization is a rare consequence of virus infection; pathogenic retroviruses are not always endogenous in their susceptible hosts, and some MLVs isolated from wild mouse populations are not endogenous, like hortulanus murine leukemia virus (HoMLV) from eastern European mice (57) and amphotropic MLV in California wild mice (58).

Alternative explanations for the observed mutational changes that alter the receptor function of the avian XPR1s, for their positive selection, and for their species and geographic distribution include the possibility that other pathogens have converged on XPR1. This is reminiscent of the fact that a mutation in the HIV-1 coreceptor CCR5 is prevalent in humans in Eurasia but originated well before human exposure to HIV-1 (59, 60). The emergence of this CCR5 mutation and its geographic distribution are consistent with a strong selective event, like an epidemic, and although the responsible pathogen has not been identified, CCR5 is the receptor for a toxin produced by *Staphylococcus aureus* (61). At this point, however, no infectious agents other than XP-MLVs are known to interact with XPR1. A second alternative explanation for the species distribution of these mutations is that the *XPR1* mutations may have arisen in ancestral Galliformes or are shared due to possible interbreeding between *Gallus* species. However, examination of the pattern of substitutions in *XPR1* exon and intron sequences does not suggest any recent recombination events, and analysis using the GARD program (reference 62 and data not shown) finds only weak evidence of recombination at a single site following codon 473 in ECL3 ($P = 0.05$) that comes from outside the galliform clade and therefore does not account for the presence of the two receptor-disabling mutations in chickens.

Domestication has a major impact on emerging diseases, and the antiviral *XPR1* genotypes with the double ECL3/ECL4 mutations are found exclusively in domestic chickens and domesticated (laboratory) mice. For the mouse, domestication was begun by Asian hobbyists who produced the fancy mouse progenitors of inbred laboratory strains (63). These hobbyists inadvertently interbred subspecies with active XP-MLVs with subspecies carrying the permissive XPR1 receptor. Domestication also creates high-risk environments with increased opportunity for jumps between and within species, and gammaretroviruses can be transmitted horizontally, as has been shown for wild and laboratory mice (64, 65). The fact that the highly restrictive receptor of domesticated chickens is not also found in domesticated turkeys and ducks is not surprising, as turkeys evolved and were domesticated in North America and therefore were not exposed to X-MLV-infected mice. The virus-susceptible duck is a waterfowl with a wide distribution that was likely domesticated in China about 2,000 years ago from mallard progenitors indigenous to an area where *M. m. musculus* likely predominates (44, 45, 66).

Virus-resistant XPR1 receptors are produced by a small number of rare mutations at the same two sites in birds (K496Q/E and Q576E) and in laboratory mice (K500E and T582Δ). One of these two sites is in an ECL (ECL4) that shows little homology between mice and birds (38%) and is also highly variable among mammals; there are numerous replacement mutations at and around the critical ECL4 site in mammals, most of which are compatible with receptor function (7) (Fig. 4). In contrast to mammalian XPR1, the avian XPR1 ECL4 is highly constrained, and positive selection was detected in the ECL4 of rodents, but not birds (6). On the other hand, for ECL3, the critical site for receptor function, K496/500, is under positive selection in birds, but not rodents. This codon spans the exon 11 consensus splice acceptor site, which imposes a bias on sequence variation for the second and third codon positions, although this bias does not disallow any of the 4 bases at the two sites (67). The restricted species and geographical distributions of these separate receptor-disabling mutations, the fact that species with restrictive receptors have been exposed to X-MLVs, the links to diet and domestication, and the observed positive selection all argue that the mutations did not occur coincidentally or through some alternative, undefined selection pressure but are the result of genetic conflicts characteristic of antiviral activity.

The observation that the laboratory mouse and various bird species in different lineages have evolved receptor-mediated resistance to X-MLVs raises the possibility that X-MLVs may be pathogenic in these species. Naturally occurring infectious and endogenous MLVs of laboratory mice and wild mice are linked to a variety of neoplastic, immunological, and neurological diseases (64, 68). In the XP-MLV family, recombinant P-MLVs derived from naturally occurring or laboratory MLV strains are pathogenic in mice (69), but there is limited work on the pathogenic potential of X-MLVs in mice because the belief that all mice are resistant to these viruses was only recently dispelled (70, 71). While it is not yet known if exposure of birds to X-MLVs results in infection and if these viruses are pathogenic in birds, our observations that duck cells support replication of infectious virus indicates that there are no significant postentry host blocks to replication. Further assessment of the possible adaptive advantage of the mutations described here should contribute to our knowledge of the host-pathogen interactions affecting transspecies transmissions in natural populations and in laboratory and agricultural stocks and should help define the evolutionary pathways that thwart those transmissions.

ACKNOWLEDGMENTS

This work was supported by the Intramural Research Program of the NIAID, NIH.

We thank Qingping Liu and Ronald Plishka for technical assistance and Devinka Bamunusinghe for help preparing figures. We also thank Rebecca Kimball (University of Florida, Gainesville, FL), James Dean (National Museum of Natural History, Washington, DC), and Glenn Proudfoot (Vassar College, Poughkeepsie, NY) for avian DNAs and tissue samples.

REFERENCES

1. **Cleaveland S, Haydon DT, Taylor L.** 2007. Overviews of pathogen emergence: which pathogens emerge, when and why? Curr. Top. Microbiol. Immunol. **315:**85–111.
2. **Levy JA.** 1973. Xenotropic viruses: murine leukemia viruses associated with NIH Swiss, NZB, and other mouse strains. Science **182:**1151–1153.
3. **Tailor CS, Nouri A, Lee CG, Kozak C, Kabat D.** 1999. Cloning and characterization of a cell surface receptor for xenotropic and polytropic murine leukemia viruses. Proc. Natl. Acad. Sci. U. S. A. **96:**927–932.
4. **Battini J-L, Rasko JEJ, Miller AD.** 1999. A human cell-surface receptor for xenotropic and polytropic murine leukemia viruses: possible role in G protein-coupled signal transduction. Proc. Natl. Acad. Sci. U. S. A. **96:**1385–1390.

5. **Yang Y-L, Guo L, Xu S, Holland CA, Kitamura T, Hunter K, Cunningham JM.** 1999. Receptors for polytropic and xenotropic mouse leukaemia viruses encoded by a single gene at *Rmc1*. Nat. Genet. **21:**216–219.
6. **Yan Y, Liu Q, Wollenberg K, Martin C, Buckler-White A, Kozak CA.** 2010. Evolution of functional and sequence variants of the mammalian XPR1 receptor for mouse xenotropic gammaretroviruses and the human-derived XMRV. J. Virol. **84:**11970–11980.
7. **Kozak CA.** 2010. The mouse "xenotropic" gammaretroviruses and their XPR1 receptor. Retrovirology **7:**101.
8. **Marin M, Tailor CS, Nouri A, Kozak SL, Kabat D.** 1999. Polymorphisms of the cell surface receptor control mouse susceptibilities to xenotropic and polytropic leukemia viruses. J. Virol. **73:**9362–9368.
9. **Yan Y, Liu Q, Kozak CA.** 2009. Six host range variants of the xenotropic/polytropic gammaretroviruses define determinants for entry in the XPR1 cell surface receptor. Retrovirology **6:**87.
10. **Fischinger PJ, Nomura S, Bolognesi DP.** 1975. A novel murine oncornavirus with dual eco- and xenotropic properties. Proc. Natl. Acad. Sci. U. S. A. **72:**5150–5155.
11. **Hartley JW, Wolford NK, Old LJ, Rowe WP.** 1977. New class of murine leukemia-virus associated with development of spontaneous lymphomas. Proc. Natl. Acad. Sci. U. S. A. **74:**789–792.
12. **Dong B, Kim S, Hong S, Das Gupta J, Malathi K, Klein EA, Ganem D, DeRisi JL, Chow SA, Silverman RH.** 2007. An infectious retrovirus susceptible to an IFN antiviral pathway from human prostate tumors. Proc. Natl. Acad. Sci. U. S. A. **104:**1655–1660.
13. **Kozak CA, O'Neill RR.** 1987. Diverse wild mouse origins of xenotropic, mink cell focus-forming, and two types of ecotropic proviral genes. J. Virol. **61:**3082–3088.
14. **Levy JA.** 1975. Host range of murine xenotropic virus: replication in avian cells. Nature **253:**140–142.
15. **Oie HK, Russell EK, Dotson JH, Rhoads JM, Gazdar AF.** 1976. Host range properties of murine xenotropic and ecotropic type-C viruses. J. Natl. Cancer Inst. **56:**423–426.
16. **Frisby DP, Weiss RA, Roussel M, Stehelin D.** 1979. The distribution of endogenous chicken retrovirus sequences in the DNA of galliform birds does not coincide with avian phylogenetic relationships. Cell **17:**623–634.
17. **Martin J, Herniou E, Cook J, O'Neill RW, Tristem M.** 1999. Interclass transmission and phyletic host tracking in murine leukemia virus-related retroviruses. J. Virol. **73:**2442–2449.
18. **Yan Y, Knoper RC, Kozak CA.** 2007. Wild mouse variants of envelope genes of xenotropic/polytropic mouse gammaretroviruses and their XPR1 receptors elucidate receptor determinants of virus entry. J. Virol. **81:**10550–10557.
19. **Cloyd MW, Thompson MM, Hartley JW.** 1985. Host range of mink cell focus-inducing viruses. Virology **140:**239–248.
20. **Peebles PT.** 1975. An *in vitro* focus-induction assay for xenotropic murine leukemia virus, feline leukemia virus C, and the feline-primate viruses RD-114/CCC/M-7. Virology **67:**288–291.
21. **Lander MR, Chattopadhyay SK.** 1984. A *Mus dunni* cell line that lacks sequences closely related to endogenous murine leukemia viruses and can be infected by ecotropic, amphotropic, xenotropic, and mink cell focus-forming viruses. J. Virol. **52:**695–698.
22. **Gillin FD, Roufa DJ, Beaudet AL, Caskey CT.** 1972. 8-Azaguanine resistance in mammalian cells. I. Hypoxanthine-guanine phosphoribosyltransferase. Genetics **72:**239–252.
23. **International Chicken Genome Sequencing Consortium.** 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature **432:**695–716.
24. **Miller DG, Miller AD.** 1992. Tunicamycin treatment of CHO cells abrogates multiple blocks to retrovirus infection, one of which is due to a secreted inhibitor. J. Virol. **66:**78–84.
25. **Hackett SJ, Kimball RT, Reddy S, Bowie RC, Braun EL, Braun MJ, Chojnowski JL, Cox WA, Han KL, Harshman J, Huddleston CJ, Marks BD, Miglia KJ, Moore WS, Sheldon FH, Steadman DW, Witt CC, Yuri T.** 2008. A phylogenomic study of birds reveals their evolutionary history. Science **320:**1763–1768.
26. **Kimball RT, Mary CM, Braun EL.** 2011. A macroevolutionary perspective on multiple sexual traits in the phasianidae (galliformes). Int. J. Evol. Biol. **2011:**423938. doi:10.4061/2011/423938.
27. **Fumihito A, Miyake T, Takada M, Shingu R, Endo T, Gojobori T, Kondo N, Ohno S.** 1996. Monophyletic origin and unique dispersal patterns of domestic fowls. Proc. Natl. Acad. Sci. U. S. A. **93:**6792–6795.
28. **Ericson PG, Anderson CL, Britton T, Elzanowski A, Johansson US, Kallersjo M, Ohlson JI, Parsons TJ, Zuccon D, Mayr G.** 2006. Diversification of Neoaves: integration of molecular sequence data and fossils. Biol. Lett. **2:**543–547.
29. **Yang Z.** 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. **13:**555–556.
30. **Dalloul RA, Long JA, Zimin AV, Aslam L, Beal K, Blomberg LA, Bouffard P, Burt DW, Crasta O, Crooijmans RP, Cooper K, Coulombe RA, De S, Delany ME, Dodgson JB, Dong JJ, Evans C, Frederickson KM, Flicek P, Florea L, Folkerts O, Groenen MA, Harkins TT, Herrero J, Hoffmann S, Megens HJ, Jiang A, de Jong P, Kaiser P, Kim H, Kim KW, Kim S, Langenberger D, Lee MK, Lee T, Mane S, Marcais G, Marz M, McElroy AP, Modise T, Nefedov M, Notredame C, Paton IR, Payne WS, Pertea G, Prickett D, Puiu D, Qioa D, Raineri E, Ruffier M, Salzberg SL, Schatz MC, Scheuring C, Schmidt CJ, Schroeder S, Searle SM, Smith EJ, Smith J, Sonstegard TS, Stadler PF, Tafer H, Tu ZJ, Van Tassell CP, Vilella AJ, Williams KP, Yorke JA, Zhang L, Zhang HB, Zhang X, Zhang Y, Reed KM.** 2010. Multi-platform next-generation sequencing of the domestic turkey (*Meleagris gallopavo*): genome assembly and analysis. PLoS Biol. **8:**e1000475. doi:10.1371/journal.pbio.1000475.
31. **Warren WC, Clayton DF, Ellegren H, Arnold AP, Hillier LW, Kunstner A, Searle S, White S, Vilella AJ, Fairley S, Heger A, Kong L, Ponting CP, Jarvis ED, Mello CV, Minx P, Lovell P, Velho TA, Ferris M, Balakrishnan CN, Sinha S, Blatti C, London SE, Li Y, Lin YC, George J, Sweedler J, Southey B, Gunaratne P, Watson M, Nam K, Backstrom N, Smeds L, Nabholz B, Itoh Y, Whitney O, Pfenning AR, Howard J, Volker M, Skinner BM, Griffin DK, Ye L, McLaren WM, Flicek P, Quesada V, Velasco G, Lopez-Otin C, Puente XS, Olender T, Lancet D, Smit AF, Hubley R, Konkel MK, Walker JA, Batzer MA, Gu W, Pollock DD, Chen L, Cheng Z, Eichler EE, Stapley J, Slate J, Ekblom R, Birkhead T, Burke T, Burt D, Scharff C, Adam I, Richard H, Sultan M, Soldatov A, Lehrach H, Edwards SV, Yang SP, Li X, Graves T, Fulton L, Nelson J, Chinwalla A, Hou S, Mardis ER, Wilson RK.** 2010. The genome of a songbird. Nature **464:**757–762.
32. **Sawai H, Kim HL, Kuno K, Suzuki S, Gotoh H, Takada M, Takahata N, Satta Y, Akishinonomiya F.** 2010. The origin and genetic variation of domestic chickens with special reference to junglefowls *Gallus g. gallus* and *G. varius*. PLoS One **5:**e10639. doi:10.1371/journal.pone.0010639.
33. **Sawyer SL, Wu LI, Emerman M, Malik HS.** 2005. Positive selection of primate *TRIM5*α identifies a critical species-specific retroviral restriction domain. Proc. Natl. Acad. Sci. U. S. A. **102:**2832–2837.
34. **Yan Y, Buckler-White A, Wollenberg K, Kozak CA.** 2009. Origin, antiviral function and evidence for positive selection of the gammaretrovirus restriction gene *Fv1* in the genus *Mus*. Proc. Natl. Acad. Sci. U. S. A. **106:**3259–3263.
35. **Sanville B, Dolan MA, Wollenberg K, Yan Y, Martin C, Yeung ML, Strebel K, Buckler-White A, Kozak CA.** 2010. Adaptive evolution of *Mus Apobec3* includes retroviral insertion and positive selection at two clusters of residues flanking the substrate groove. PLoS Pathog. **6:**e1000974. doi:10.1371/journal.ppat.1000974.
36. **Denner J.** 2007. Transspecies transmissions of retroviruses: new cases. Virology **369:**229–233.
37. **Tarlinton RE, Meers J, Young PR.** 2006. Retroviral invasion of the koala genome. Nature **442:**79–81.
38. **Adler J, Lawler A.** 2012. How the chicken conquered the world. Smithsonian **43:**40–47.
39. **Liu YP, Wu GS, Yao YG, Miao YW, Luikart G, Baig M, Beja-Pereira A, Ding ZL, Palanichamy MG, Zhang YP.** 2006. Multiple maternal origins of chickens: out of the Asian jungles. Mol. Phylogenet. Evol. **38:**12–19.
40. **West B, Zhou BX.** 1989. Did chickens go north—new evidence for domestication. W. Poult. Sci. J. **45:**205–218.
41. **Kanginakudru S, Metta M, Jakati RD, Nagaraju J.** 2008. Genetic evidence from Indian red jungle fowl corroborates multiple domestication of modern day chicken. BMC Evol. Biol. **8:**174.
42. **Darwin C.** 1868. The variation of animals and plants under domestication, authorized ed. O. Judd & Company, New York, NY.
43. **Eriksson J, Larson G, Gunnarsson U, Bed'hom B, Tixier-Boichard M, Stromstedt L, Wright D, Jungerius A, Vereijken A, Randi E, Jensen P, Andersson L.** 2008. Identification of the *yellow skin* gene reveals a hybrid origin of the domestic chicken. PLoS Genet. **4:**e1000010. doi:10.1371/journal.pgen.1000010.
44. **Boursot P, Auffray JC, Britton-Davidian J, Bonhomme F.** 1993. The evolution of house mice. Annu. Rev. Ecol. Syst. **24:**119–152.

45. **Marshall J.** 1981. Taxonomy, p 17–26. *In* Foster HL, Small JD, Fox JG (ed), The mouse in biomedical research, vol 1. Academic Press, New York, NY.
46. **Sage R, Atchley WR, Capanna E.** 1993. House mice as models in systematic biology. Syst. Biol. **42:**523–561.
47. **Maryanto I, Kitchener DJ, Prijono SN.** 2005. Morphological analysis of house mice, *Mus musculus* (Rodentia, Muridae) in Southern and Eastern Indonesia and Western Australia. Mamm. Study **30:**53–63.
48. **Yonekawa H, Moriwaki K, Gotoh O, Miyashita N, Matsushima Y, Shi L, Cho W, Zhen X, Tagashira Y.** 1988. Hybrid origin of Japanese mice "*Mus musculus molossinus*": evidence from restriction analysis of mitochondrial DNA. Mol. Biol. Evol. **5:**63–78.
49. **Orth A, Adama T, Din W, Bonhomme F.** 1998. Natural hybridization between two subspecies of the house mouse, *Mus musculus domesticus* and *Mus musculus castaneus*, near Lake Casitas, California. Genome **41:**104–110. (In French.)
50. **Yang H, Wang JR, Didion JP, Buus RJ, Bell TA, Welsh CE, Bonhomme F, Yu AH, Nachman MW, Pialek J, Tucker P, Boursot P, McMillan L, Churchill GA, de Villena FP.** 2011. Subspecific origin and haplotype diversity in the laboratory mouse. Nat. Genet. **43:**648–655.
51. **Chattopadhyay SK, Oliff AI, Linemeyer DL, Lander MR, Lowy DR.** 1981. Genomes of murine leukemia viruses isolated from wild mice. J. Virol. **39:**777–791.
52. **Kozak CA, Hartley JW, Morse HC, III.** 1984. Laboratory and wild-derived mice with multiple loci for production of xenotropic murine leukemia virus. J. Virol. **51:**77–80.
53. **Lieber M, Sherr C, Potter M, Todaro G.** 1975. Isolation of type-C viruses from the Asian feral mouse *Mus musculus molossinus*. Int. J. Cancer **15:**211–220.
54. **Morse HC, III, Kozak CA, Yetter RA, Hartley JW.** 1982. Unique features of retrovirus expression in F/St mice. J. Virol. **43:**1–7.
55. **East J, Tilly RJ, Tuffrey M, Harvey JJ.** 1978. The early appearance and subsequent distribution of murine leukaemia virus in NZB embryos. Int. J. Cancer **22:**495–502.
56. **Bolisetty M, Blomberg J, Benachenhou F, Sperber G, Beemon K.**2012. Unexpected diversity and expression of avian endogenous retroviruses. mBio **3:**e00344–00312. doi:10.1128/mBio.00344-12.
57. **Voytek P, Kozak CA.** 1989. Nucleotide sequence and mode of transmission of the wild mouse ecotropic virus, HoMuLV. Virology **173:**58–67.
58. **O'Neill RR, Hartley JW, Repaske R, Kozak CA.** 1987. Amphotropic proviral envelope sequences are absent from the *Mus* germ line. J. Virol. **61:**2225–2231.
59. **Stephens JC, Reich DE, Goldstein DB, Shin HD, Smith MW, Carrington M, Winkler C, Huttley GA, Allikmets R, Schriml L, Gerrard B, Malasky M, Ramos MD, Morlot S, Tzetis M, Oddoux C, di Giovine FS, Nasioulas G, Chandler D, Aseev M, Hanson M, Kalaydjieva L, Glavac D, Gasparini P, Kanavakis E, Claustres M, Kambouris M, Ostrer H, Duff G, Baranov V, Sibul H, Metspalu A, Goldman D, Martin N, Duffy D, Schmidtke J, Estivill X, O'Brien SJ, Dean M.** 1998. Dating the origin of the *CCR5-Δ32* AIDS-resistance allele by the coalescence of haplotypes. Am. J. Hum. Genet. **62:**1507–1515.
60. **Libert F, Cochaux P, Beckman G, Samson M, Aksenova M, Cao A, Czeizel A, Claustres M, de la Rua C, Ferrari M, Ferrec C, Glover G, Grinde B, Guran S, Kucinskas V, Lavinha J, Mercier B, Ogur G, Peltonen L, Rosatelli C, Schwartz M, Spitsyn V, Timar L, Beckman L, Parmentier M, Vassart G.** 1998. The Δ*ccr5* mutation conferring protection against HIV-1 in Caucasian populations has a single and recent origin in Northeastern Europe. Hum. Mol. Genet. **7:**399–406.
61. **Alonzo F, III, Kozhaya L, Rawlings SA, Reyes-Robles T, DuMont AL, Myszka DG, Landau NR, Unutmaz D, Torres VJ.** 2013. CCR5 is a receptor for Staphylococcus aureus leukotoxin ED. Nature **493:**51–55.
62. **Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SD.** 2006. Automated phylogenetic detection of recombination using a genetic algorithm. Mol. Biol. Evol. **23:**1891–1901.
63. **Morse HC, III.** 1978. Introduction, p 1–31. *In* Morse HC, III (ed), Origins of inbred mice. Academic Press, New York, NY.
64. **Gardner MB, Chiri A, Dougherty MF, Casagrande J, Estes JD.** 1979. Congenital transmission of murine leukemia virus from wild mice prone to the development of lymphoma and paralysis. J. Natl. Cancer Inst. **62:**63–70.
65. **Portis JL, McAtee FJ, Hayes SF.** 1987. Horizontal transmission of murine retroviruses. J. Virol. **61:**1037–1044.
66. **Qu L, Liu W, Yang F, Hou Z, Zheng J, Xu G, Yang N.** 2009. Origin and domestication history of Peking ducks determined through microsatellite and mitochondrial marker analysis. Sci. China Life Sci. **52:**1030–1035.
67. **Mount SM.** 1982. A catalogue of splice junction sequences. Nucleic Acids Res. **10:**459–472.
68. **Dudley JP, Mertz JA, Bhadra S, Palmarini M, Kozak CA.** 2010. Endogenous retroviruses and cancer, p 119–162. *In* Dudley J. (ed), Retroviruses and insights into cancer. Springer, New York, NY.
69. **Cloyd MW, Hartley JW, Rowe WP.** 1980. Lymphomagenicity of recombinant mink cell focus-inducing murine leukemia viruses. J. Exp. Med. **151:**542–552.
70. **Baliji S, Liu Q, Kozak CA.** 2010. Common inbred strains of the laboratory mouse that are susceptible to infection by mouse xenotropic gammaretroviruses and the human-derived retrovirus XMRV. J. Virol. **84:**12841–12849.
71. **Sakuma T, Tonne JM, Squillace KA, Ohmine S, Thatava T, Peng KW, Barry MA, Ikeda Y.** 2011. Early events in retrovirus XMRV infection of the wild-derived mouse *Mus pahari*. J. Virol. **85:**1205–1213.